Responses to Review #2

The author would like to thank the reviewer for his valuable comments which helped in improving the quality of the manuscript. My point-by-point responses to the reviewer's comments appear in bold below.

Major comments

Robustness and uncertainty of aerosol classification

The aerosol classification is based on the joint use of LR and PDR (Section 4.2, Table 3). While this approach is well established, it relies on predefined ranges of LR and PDR for different aerosol types.

However:

- The ranges assigned to different aerosol classes (e.g. dust vs pollution) partly overlap.
- The classification is applied deterministically without providing any estimate of uncertainty or confidence.

Given the frequent occurrence of mixed aerosol layers (as acknowledged in the manuscript), this may lead to ambiguous classification.

Suggestion:

Provide a discussion on the uncertainty and potential misclassification of aerosol types.

If possible, include a probabilistic interpretation or at least comment on the sensitivity of the classification to the assumed LR/PDR ranges.

The LR and PDR classes in Table 3 were estimated using data from various lidar observation campaigns. These campaigns reveal considerable variability in the values for the different aerosol types. This also depends on whether mixing occurred during the measurements. I have expanded the justification for considering the LR and PDR values for the three aerosol classes. These values are highlighted in independent measurement campaigns.

The text in Section 4.2 has therefore been updated as follows:

"…

Dust aerosols over the western Mediterranean Sea are attributable to the long-range transport of Saharan dust. These aerosols are typically associated with LRs between 30 and 80 sr (Papayannis et al., 2008), specifically for high PDRs, above 10% and up to ~30%. This variability is explained by the level of mixing with other aerosol types and the nature of soils in uplift areas. In Mallet et al. (2022), the reported LR values at 355 nm are predominantly between 32 and 84 sr, including polluted dust aerosols. In Groß et al. (2015), however, they are closer to 55±20 sr, and for the same authors, the PDRs are of the order of 25±15%. These values include those provided by Chazette and Totems (2023), which range from 45 to 70 sr for the LR and from 20±15% for the PDR.

Marine aerosols have been found to have low LRs, ranging from 20 to 30 sr (Flamant et al., 1998), which corresponds to a very low PDR. However, these particles are often mixed

with polluted aerosols or dust, which increases their LR and PDR. Mallet et al. (2022) reported LR of ~25±6 sr, which is very close to the values given in Groß et al. (~25±10 sr). The PDR remains low due to the spherical characteristics of these highly hydrophilic aerosols, with values of ~3±3% (e.g. Chazette et al., 2019; Groß et al., 2015).

The LR of pollution and biomass fire aerosols is highly variable depending on the combustion source, with values of ~60±20 sr (Chazette et al., 2019; Groß et al., 2015). Mallet et al. (2022) also report smoke-like aerosols with an LR between 42 and 73 sr. It should be noted that the combustion temperature influences the LR by affecting the chemical composition of the aerosols. It also influences the PDR, mainly in the event of strong thermal convection, whereby terrigenous particles can be lifted from the surface, leading to PDRs of the order of 8% (Chazette et al., 2016a). By contrast, PDRs typically range from 2 to 3% for young aerosols, rising to over 6% after long-range transport. Groß et al. (2015) reported PDRs of ~3±3% for smoke-like aerosols. These values are similar to those provided by Chazette and Totems (2023).

All of these values are obtained from short-term campaigns in very specific locations, and are few in numbers. Significant variation can clearly occur depending on the degree of particle mixing and potential ageing during transport.

Figure 4 also allows us to determine the range of variation for the LR and PDR. For dust-like aerosols, LR varies from approximately 30 to 70 sr, whereas PDR is around 20%± 10%. This is broadly consistent with previous reports. Similarly, the LR for marine-like (soluble) aerosols tends to be between 20 and 40 sr, with PDRs ranging from ~0 to 8%. Pollution-like (carbonaceous) aerosols, on the other hand, are associated with LR values between 40 and 80 sr and PDRs of up to 10%.

…"

I have also added the following after Eq. 10:

"…The exponential form clarifies potential mixtures by assigning lower weights to extreme values, while still taking the large variability associated with each type of aerosol into account…"

I believe all this provides a clearer picture of the variability of each type of aerosol, and consequently of the potential uncertainties in the classification. It also demonstrates how these factors were considered when selecting the parameters for the Gaussian distributions.

Lack of uncertainty propagation

Although uncertainties are provided for individual retrieved quantities (AEC, LR, PDR), there is no propagation of these uncertainties into: aerosol classification, comparison with CAMS, or the final conclusions. Given that the study draws quantitative conclusions based on correlation, bias, and RMSE, this is an important limitation.

Suggestion:

Include a discussion on how measurement uncertainties may impact: classification results, and the evaluation of CAMS performance. A full propagation may not be necessary, but a quantitative or semi-quantitative assessment would be valuable.

The AEC is not a factor in aerosol classification. As previously mentioned, only LR and PDR are used. A Monte Carlo approach can be used to assess how the classification changes based on the determined LR and PDR values and by increasing the relative errors to 5% for LR and the absolute errors to 2% for PDR. I carried out this study on two profiles where the three aerosol types that reached the highest altitudes were present, adding the following elements to subsection 4.2.:

"A sensitivity analysis using a Monte Carlo approach was performed to assess the reliability of this optical apportionment based on lidar measurements, as described by Royer et al. (2011). Uncertainties in LR and PDR were assumed to be up to 5% (relative) for LR and 2% (absolute) for PDR. Figure 6 presents the results for the three particle types. As can be seen, the influence of the uncertainties on lidar-derived optical parameters is small, remaining below 5% for each aerosol type. This demonstrates the robustness of the classification with respect to realistic uncertainties associated with lidar measurements."

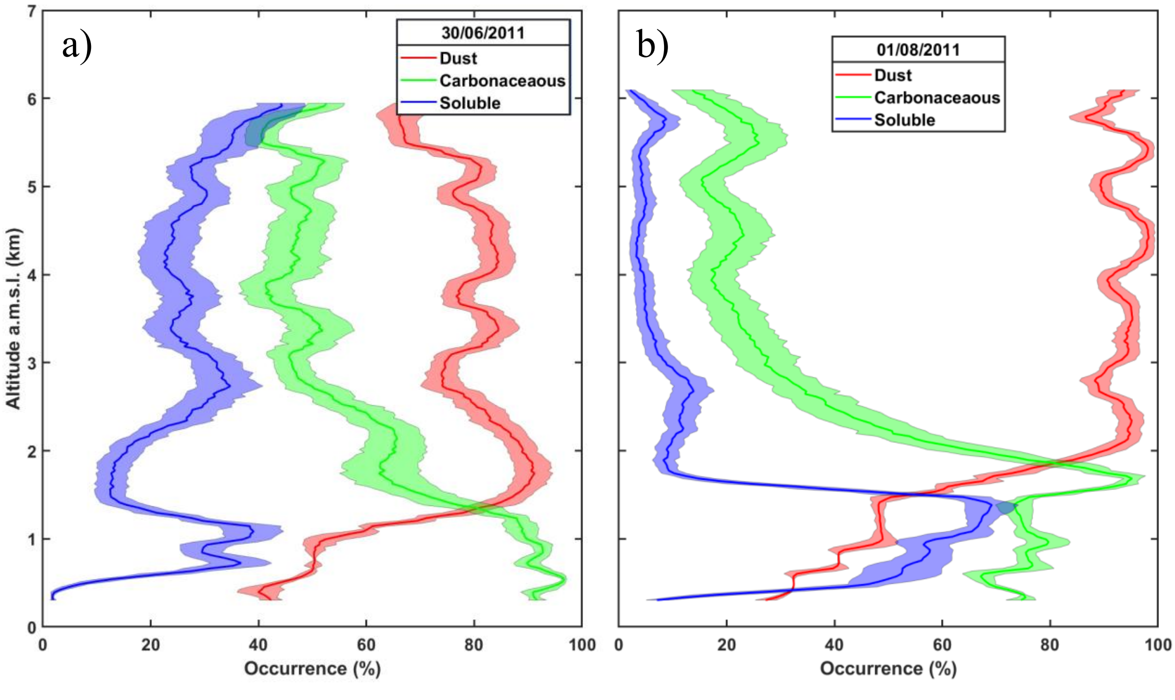
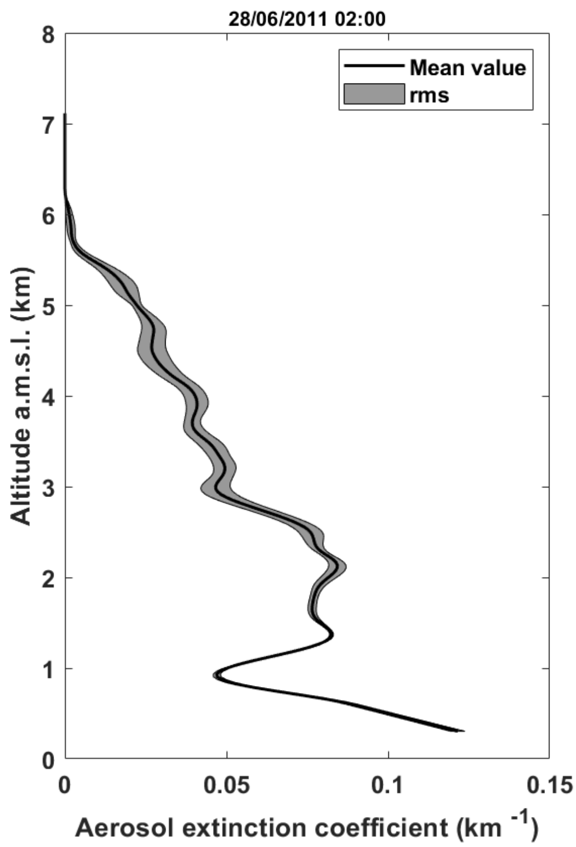


Figure 6. Uncertainty on the occurrences of each aerosol type due to statistical error on both the LR (5% in relative) and PDR (2% in absolute) on 30 June 2011 and 1 August 2011. The mean value is given by the continuous lines and the standard deviation by the shaded area.

When comparing AEC profiles, uncertainties in the lidar measurements (RMS<0.01 km^{-1} for the average profiles) have little impact on the model's performance (see the figure below for an example). Furthermore, these uncertainties are accounted for by the variability of the profiles and, consequently, in the statistical results. This is even more

evident for AOTs, which integrate AEC profiles. For example, a vertical profile of the AEC with its uncertainty (gray area) is given below:



Vertical representativeness mismatch between lidar and CAMS

The comparison between lidar profiles and CAMS data (Section 5.2) does not explicitly account for differences in vertical resolution and representativeness, as Lidar profiles have high vertical resolution (~100 m), while CAMS profiles are much coarser and model-based.

The observed discrepancies, particularly in the boundary layer (low correlation and positive bias), are attributed mainly to model deficiencies. However, part of these differences may arise from scale mismatch and representativeness errors, especially in a coastal environment.

Suggestion:

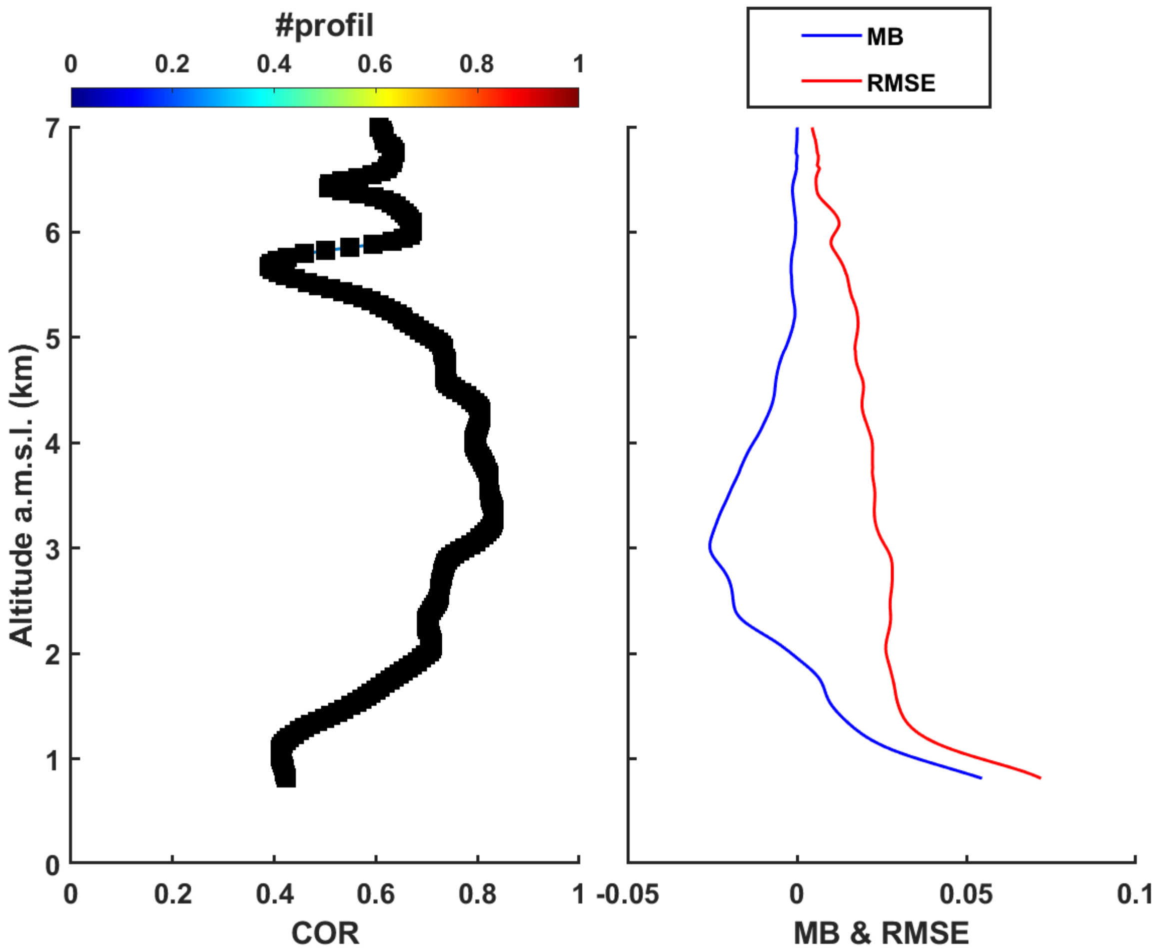
- Discuss the impact of vertical resolution mismatch.
- If possible, consider smoothing lidar profiles to the CAMS vertical grid, or explicitly acknowledge this limitation.

There is indeed a significant difference in vertical resolution between the lidar measurements and the CAMS data. This is one of the challenges involved in comparing vertical profiles. In this article, this challenge was overcome by down sampling the lidar data to match the resolution of the CAMS profiles using spline interpolation. I verified that similar results could be obtained using the lidar's vertical resolution, again via spline

interpolation applied to model data. The result is shown below, but I have not included it in the article as it adds nothing further. I have added the following text to subsection 5.2:

"The significant difference in resolution between the CAMS and LiDAR data requires the application of spline interpolation (Perperoglou et al., 2019). The results presented here are at the model resolution. Similar results are obtained when the vertical resolution of the lidar is used with spline interpolation applied to the model data. It is worth noting that the model's resolution is higher for the lower layers, meaning they are estimated more accurately. However, for the upper layers, the thickness of the desert dust clouds limits the effectiveness of the model's vertical resolution."

Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., and Schmid, M.: A review of spline function procedures in R, https://doi.org/10.1186/s12874-019-0666-3, 6 March 2019.



Interpretation of CAMS performance

The manuscript concludes that CAMS shows "excellent agreement" with lidar observations. While this is valid for the free troposphere, the results clearly show significantly lower correlation in the boundary layer (~0.5) and systematic bias near the surface. Therefore, the conclusions appear somewhat over-generalized.

Suggestion:

Refine the conclusions to clearly distinguish between:

- good performance in the free troposphere,
- and more limited performance in the boundary layer.

I have separated the two layers and strengthened the conclusion:

"…In the free troposphere, the CAMS reanalyses (EAC4) were found to be in very good agreement with the geophysical products derived from optical lidar measurements for dust layers. This is particularly notable for the AOTs. The reanalyses accurately reproduce dust events in terms of both time and vertical extends with COR ~0.87 with a slight overestimation of the AOT (0.01).
In the lower troposphere, below 2 km a.m.s.l., the mixing of pollution (characterised by the carbonaceous component) with marine (characterised by the soluble component) and dust contributions is also accurately reproduced throughout the two-month experiment. Nevertheless, the correlation observed between lidar measurements and CAMS reanalyses is significantly lower (0.56). It is associated with an underestimation of the contribution to the AOT of this layer of around 0.06 at 355 nm, which may be due to an incorrect assessment of emission processes. Nevertheless, it should be noted that failure to properly consider the hygroscopicity of aerosols may affect this result. Such an oversight may be present in the model itself, given that a number of assumptions have been made. However, it may also arise from the assessment of aerosol extinction coefficients performed in this study. Note that, on the other hand, lidar measurements take into account the hydrophilic nature of aerosols.

In the entire atmospheric column, an agreement between 10 and 20% on the temporal evolution of the total AOTs derived from CAMS and lidar is highlighted. This is associated with a small mean bias (-0.01). This is primarily due to the fact that dust aerosols dominated the aerosol load during the FENNEC field campaign and the reanalyses better represent them…"

Assumptions in CAMS extinction reconstruction

The method used to reconstruct extinction profiles from CAMS mixing ratios (Section 3.2) assumes constant specific cross-sections for each aerosol type. This simplification neglects hygroscopic growth (especially relevant near the surface), variability in aerosol composition within each class, and size distribution changes. These factors may contribute significantly to the discrepancies observed in the lower troposphere.

Suggestion:

- Expand the discussion of these assumptions and their impact.
- If possible, comment on the sensitivity of the results to these choices.

The answer was provided in detail in the replies to reviewer #1, so I will not repeat it here. A discussion was indeed lacking. It should be noted that the calculation based on the model data is justified not only by the meteorological conditions, but also by an understanding of the model's underlying assumptions. For instance, bin size does not change due to

growth with humidity, and it is assumed that sea salt has 80% relative humidity (RH). Table 2 of the article now illustrates the impact of the calculation assumptions on the AOTs for each type of aerosol. The effect remains small, though it is slightly higher for soluble aerosols in terms of correlation (0.82).

When comparing with lidar data, it is important to bear in mind that the measurement considers the aerosol as a whole and therefore includes potential mixtures. No assumptions are made during inversion of the profiles. The statistical results therefore provide an incomplete representation of the model's estimation of the effect of relative humidity (RH). Nevertheless, according to the RH values derived from ERA5, it should be noted that the deliquescence points are not necessarily reached. Having compared ERA5 profiles with radiosonde and lidar observations, I have considerable confidence in these reanalyses.

Temporal averaging strategy

The lidar profiles are averaged over relatively long nighttime periods. While this is necessary for signal-to-noise reasons, it may smooth out short-term variability, and potentially improve apparent agreement with model data.

Suggestion:

- Include a discussion of the implications of this temporal averaging.
- Optionally, illustrate one case with higher temporal resolution.

Novelty and relevance of the dataset

The dataset dates from 2011, and the manuscript does not sufficiently justify its relevance in the context of more recent observations and advances in instrumentation and modelling.

This point was also raised by Referee #1.

Suggestion:

- Clearly state the added value of this dataset

The relevance of the dataset has been clarified.

Although the dataset is old, it remains entirely relevant. It is the longest continuous dataset available from our lidars. Furthermore, I can confidently state that lidar technology has not changed significantly since 2011. LAASURS is a French lidar system that remains operational and uses proven technology employed in recent developments and field campaigns (Laly et al., 2024). The FENNEC dataset has the additional advantage of having been collected during a period of low cloud cover in southern Spain. This provided sufficient nights for meaningful statistical analysis, following a robust inversion process.

I explained this in subsection 2.1 and 2.2.2.

"It is worth noting that the FENNEC ground-based experiment follows the Mediterranean Dust Experiment (MEDUSE) (Hamonou et al., 1999; Dulac and Chazette, 2003) and predates the Chemistry-Aerosol Mediterranean Experiment (ChArMEx),

conducted between 2013 and 2014 (Chazette et al., 2016; Di Girolamo et al., 2020; Chazette et al., 2019; Mallet et al., 2015). The latter notably included contributions from the European Aerosol Research Lidar Network (EARLINET) (Barragan et al., 2017; Navas-Guzmán et al., 2013). Compared with these campaigns, FENNEC provided an exceptionally long and continuous lidar sampling of the troposphere in cloud-free conditions. "

"LAASURS (Erreur ! Source du renvoi introuvable.**b) itself is described in detail and validated in Royer et al. (2011a) and Chazette et al. (2019b) and it has been used in numerous field campaigns, such as the journey from Paris to Lake Baikal (Dieudonné et al., 2015). LAASURS is a French lidar system that remains operational and uses proven technology that has been employed in recent developments and field campaigns (Laly et al., 2024). "**

Link to radiative implications

The introduction emphasizes aerosol radiative effects and climate relevance, but the study does not directly quantify radiative forcing.

Suggestion:

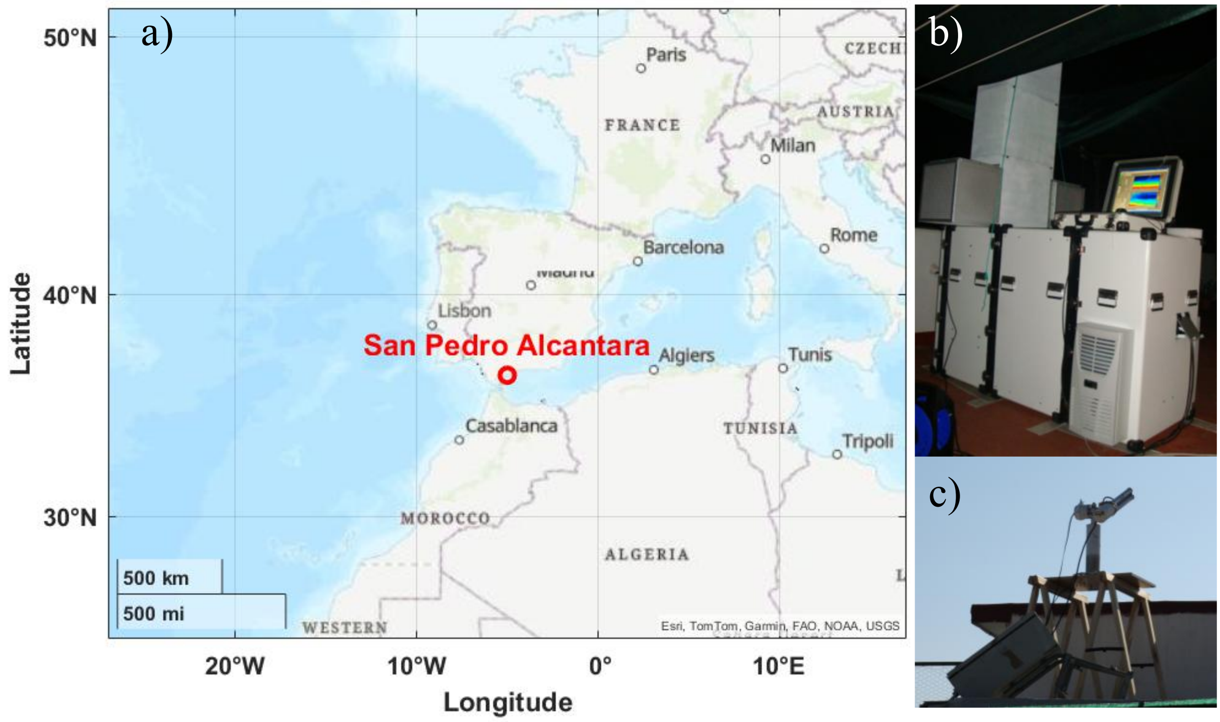
- Either briefly discuss the radiative implications of the observed aerosol properties or moderate the claims in the introduction

The introduction was too general and exceeded the scope of the article. It has now been revised to be more focused (see the responses to Review_RC1).

Minor comments

- Figure 1- the map is blurry; please improve the quality in the final uploaded file

  I think it is related to the conversion to PDF. I hope that Figure 1 below is not blurry.



- Figure 4 and 5- consider improving readability (the graphs are blurry).

  Unfortunately, it seems that converting the figures to PDF has lowered their quality. My apologies for this. Please note that the figures are provided in JPG format for the final publication.

- Terminology such as "optical apportionment" should be clearly defined when first introduced.

  It has now been defined in the introduction: (i.e. classification of aerosols using optical measurements)

- Ensure consistency in terminology for statistical metrics (RMSE vs RMSD), in line with Referee #1.

  As I wrote in my reply to Referee #1, this mainly stems from the conventions within each scientific field. These definitions are widely used and have been employed several times in our publications. For the sake of consistency with previous works, I would prefer not to change them. Moreover, since they are clearly defined in the Annex, I do not believe it will impair comprehension.

- page 12 Lines 11 and 12 "may lead to LRs of the order of 8%"; LR is measured in sr, not %. Please clarify in here what 8% (and 2-3% on the next line) mean

  Yes, the correction has been done.

- page 16 Line 10: "Qualitatively, The altitude and time locations…"; "The" should be lowercase.

  The correction has been done.

- page 18 Line 2 and 3 "rectangle" Should be corrected with " triangle"

The correction has been done.

General language issues

The manuscript would benefit from a careful language revision, as it contains a number of minor grammatical errors, typographical issues, and occasional imprecise phrasing. While these do not affect the scientific content, improving readability would significantly enhance the clarity of the paper.

Some long sentences could be split for clarity.

The manuscript has been proofread and the errors corrected.